



SHANE GREENSTEIN

DANIEL YUE

KERRY HERMAN

SARAH GULICK

Hugging Face: Serving AI on a Platform

Clément Delangue (CEO) and co-founders Julien Chaumond (Chief Technology Officer, or CTO) and Thomas Wolf (Chief Science Officer, or CSO) founded Hugging Face in 2016. The company evolved to become a platform that allowed anyone to discover and access artificial intelligence (AI) models for natural language processing (NLP)—computer programs that could complete language-based tasks such as translations or text creation. By spring 2021, it had expanded into other machine learning (ML) tasks like vision and audio. By the fall of 2022, the website, which hosted open-source ML technology, had grown to support over 100,000 daily users and 9 million downloads per day.

The technologies on Hugging Face's platform were mostly created by volunteers from its open-source community, who contributed for free, usually developing new ML models. These volunteers were motivated by goals such as accelerating the development of AI technology, honing their skills on the technical frontier, and helping others. It was all part of a broader open-source ethos prevalent in the ML community that emphasized collaboration, transparency, and teamwork. These volunteers tended to have high-level expertise across a breadth of ML applications, enabling Hugging Face to provide state-of-the-art models that could function across a diverse range of use-cases—a key aspect of the company's model. Hugging Face had little power to direct its contributors; instead, the firm adapted its development goals to its open-source community.

Given this distributed, community-driven approach to technology development, Hugging Face's operating structure was decentralized. While Delangue and his co-founders offered guidance on project prioritization and helped to connect related efforts among the company's employees and contributors, both employees and contributors were free to initiate the projects they worked on. This autonomy helped Hugging Face attract and keep talented developers, who prized autonomy and community impact over other job characteristics, such as the higher salaries offered by larger technology companies. The company was registered in New York City, and had office space in New York and Paris, but most of its more than 150 employees worked remotely. To increase transparency, the company prioritized asynchronous discussions on Slack channels over synchronous meetings. Meetings held via video call were recorded and posted on the company's Slack.

Professor Shane Greenstein, Doctoral Student Daniel Yue, Director Kerry Herman and Case Researcher Sarah Gulick (Case Research & Writing Group) prepared this case. It was reviewed and approved before publication by a company designate. Funding for the development of this case was provided by Harvard Business School and not by the company. HBS cases are developed solely as the basis for class discussion. Cases are not intended to serve as endorsements, sources of primary data, or illustrations of effective or ineffective management.

Copyright © 2022, 2023 President and Fellows of Harvard College. To order copies or request permission to reproduce materials, call 1-800-545-7685, write Harvard Business School Publishing, Boston, MA 02163, or go to www.hbsp.harvard.edu. This publication may not be digitized, photocopied, or otherwise reproduced, posted, or transmitted, without the permission of Harvard Business School.

While Hugging Face operated in many ways like an open-source, research-oriented community, it was a for-profit company backed by venture capital (VC) funding. The company began offering paid services in 2021, leaving its core services free for general use. A small percentage of users—mostly companies that used ML technologies—paid for additional ML operations (MLOps) services, enabling them to employ the Hugging Face platform for storing, training, and deploying models. Paying users could also consult internal Hugging Face experts on how to best apply ML technologies to solve their business problems. In 2022, approximately 1,000 companies paid for these services. By early fall 2022, Delangue was seeing increased interest from paying companies for certain tools—especially tabular data—but the research interests of the community did not always match the commercial interests of Hugging Face’s paying customers.

The diversity of the open-source community, which had empowered so much of Hugging Face’s success, inevitably led to unprecedented and tense situations. In late May of 2022, a controversy emerged around a new language model, GPT-4chan, which a community member had trained to recreate speech from the controversial discussion website 4chan by a community member. GPT-4chan was on the Hugging Face website for others to use, study, or refine. However, the model generated controversy within the community because it reproduced the sexist, racist, ableist, or otherwise hurtful language common to the 4chan discussion boards. Hugging Face aspired to work collaboratively with the research community developing AI, empowering them to pursue their interests. Yet, this model raised questions about whether unlimited community discretion on research topics conflicted with Hugging Face’s other principles and aspirations. After some days of internal discussion, Hugging Face decided not to support the model, but left the debate and model documentation in place for future reference. As summer turned into fall, and Hugging Face continued to grow, Delangue looked back on this situation with more reflection. He wondered if the company had made the right decision.

Industry Landscape

Hugging Face provided NLP and other ML models within a framework consistent with the practices of open-source computing. This allowed Hugging Face to continue to provide technology that remained relevant even as the technical frontier advanced.

The Growth of Machine Learning

ML was a method for creating AI in which the computer program’s output was based on algorithms learning from data by identifying patterns that could prompt inference on future patterns, rather than output based on specific rules.¹ ML could analyze numbers, textual patterns, or images, and was used to analyze data that would be too complex for human researchers to explicitly write a program about.²

Most of modern ML grew out of Geoffrey Hinton’s 1980s breakthroughs in efficiently training neural networks. Neural networks constituted a model architecture comprising sets of nodes connected via link functions that allowed the model to learn relationships in the data rather than require programmers to specify the rules upfront. This modeling approach was called deep learning, in reference to the layered sets of nodes mimicking how the human brain connected neurons.³ Hinton’s goal was to build “a bridge between computer science and biology.”⁴

The annual ImageNet competition was an important deep-learning image classification contest based on a multi-million image dataset released in 2009 by Fei-Fei Li.⁵ ImageNet comprised a large database of object images from which ML models could learn. It included over 20,000 labeled objects such as cats and trees, making it easy to train and compare models aimed at classifying those objects.⁶ Begun in 2010, the ImageNet competition quickly became the accepted community benchmark for

measuring algorithm accuracy in recognizing objects. In 2012, a massive breakthrough occurred when Hinton and his students Ilya Sutskever and Alex Krizhevsky successfully applied a new kind of neural network architecture (a convolutional neural network) to the ImageNet classification problem. That year, Hinton's team won the contest with an error rate far below 25%. They were widely imitated, and progress accelerated. A few years later, error rates were below 5%, considered the error rate for a typical human. The success of this approach began the "deep learning revolution" and laid the foundation for contemporary image recognition technology.⁷ The algorithms became so accurate that the ImageNet contest ended after 2017, and a new, more challenging contest was created in 2018. (See **Exhibit 1**.)

After this success, a number of established firms and startups began to explore commercial applications for neural networks. Within a few years, many startups appeared, targeting problems across the ML ecosystem. Their applications went beyond neural networks and vision recognition, extending to new areas, including NLP. (See **Exhibit 2** for global funding interest.)

Natural Language Processing

NLP was an approach to building models to solve problems involving natural language—that is, human communication through words. NLP technology solutions used linguistics, statistics, computer science, and AI in order to categorize and interpret natural language. The NLP field first developed structurally, with researchers describing the grammar and comprehension rules underlying language. By the 1980s, researchers had moved away from specifying rules to develop statistical methods of modeling language that computers could execute. In the 2010s, as computing power and methodologies developed further, scientists created flexible computer models that could write, understand, and translate texts at levels comparable to humans.

By the late 2010s, NLP research had gained wide interest from industry and was applied to a broad range of services. These included translation features offered by companies including Google, customer service help provided by chatbots, and automatic text generation used by marketing companies. Almost all were services that people had previously provided, and now could be automated and delivered more cheaply and on a greater scale. By the spring of 2022, many companies sought to implement AI factories for NLP by standardizing text-based datasets, improving them on existing models, and integrating these models into business processes and products. (For a diagram showing one AI industry paradigm, see **Exhibit 3**.)

Industry revenue and major players In 2021, experts estimated that the market for AI technology, of which NLP was a part, was around \$60 billion. Those most optimistic about its growth predicted that it would grow to \$422 billion by 2028.⁸

The growing industry also created demand for services that assisted businesses with AI and NLP operations, including companies that provided and cleaned data, and those that provided computing power.^a Providers tended to be major technology companies with the capacity to run such programs. Many large tech companies had separate cloud computing services, including Amazon Web Services (AWS), Google Cloud, and Microsoft Azure. Their services included AI/NLP computing power, user interfaces that made training and deploying models accessible to non-technical users, website hosting, data storage, and other Internet services. In the few years up to 2022, AWS provided at least 30% of all cloud computing services globally; Azure provided approximately 20%.⁹

^a Computing power was the use power of high-powered computers that allowed users to run models that used huge amounts of data, often millions or billions of individual pieces of information.

Open Source

Open-source software was unpriced, and, except for some occasional restrictions on commercial use, its licensing terms did not restrict who could use the software or inspect the source code. The practice of open source originated in the 1980s and formalized over time with the development of different licenses governing its sharing software for commercial and research purposes. During the spread of the Internet in the 1990s, open-source software became widespread, and technical users assembled standardized functionality from various pieces of well-known programs. In that era, for example, the acronym LAMP emerged to describe the suite of Linux, Apache, MySQL and PHP/PERL/Python – open-source software widely in use during that time.¹⁰

By nature, open source was considered collaborative and public, often worked on by many people, including users as co-developers. Open-source culture included helping new participants learn new skills, and being sure to give proper credit to each collaborator. Open-source software could often be built faster and more cheaply than programs designed by a single company, since most contributors were motivated volunteers. It also tended to be more resilient than proprietary software, given contributors' multiple perspectives.

Several open-source frameworks and languages were widely used in AI. For example, Google released Tensor Flow as open-source software in 2015, and updated it in 2019 as Tensor Flow 2.0. This software worked with many common computer languages, such as Python and JavaScript, and provided a framework for developing neural network models. Meta (formerly known as Facebook) developed another common framework, PyTorch, and released it in 2017. Many programmers preferred PyTorch's approach to developing large-scale models. Both frameworks continued to improve after their release with new tools and libraries contributed by open-source communities.

GitHub The most popular place to store and share open source programs was GitHub, founded in 2008. It provided many tools that made collaboration easier, such as a “git,” which provided version control. By 2018, GitHub had 28 million contributors. Its revenue came from premium subscriptions to private storage and other business-facing features, costing between \$7 and \$21 per user per month.¹¹

In June 2018, Microsoft acquired GitHub for \$7.5 billion.¹² For some developers using the platform, its acquisition by a major company made GitHub less trustworthy than it had been as an independent organization, although the platform remained very popular.¹³ In 2022, GitHub continued to play an important role for the open-source community. It acted as a repository for some open-source programs and libraries developed for TensorFlow and PyTorch.

Industry revenue and growth While open-source programs were released for free and were openly available for users to use and modify, many services existed to help companies adapt and sequester their open-source software for use on a private platform. For many companies, the free software saved money on licensing costs. The estimated savings were difficult to project. Due to its millions of users, researchers thought open-source Webserver software Apache alone was responsible for billions of dollars in the U.S. Internet economy.¹⁴

Many open-source software projects had become essential for the operations of organizations using them. Yet, because open-source software was free to access, it was difficult for software developers to monetize their advances. Projects that might start as a grassroots effort could then be taken on and maintained by the companies or nonprofit organizations that relied on them.

A subset of open-source software had developed “freemium” approaches. Freemium, a portmanteau of “free” and “premium,” involved providing basic services at no cost, while charging

for advanced features that enhanced the software's functionality and reliability, typically for use within large enterprises. Sometimes the charges arose when enterprises wanted advanced tools, and sometimes when enterprises used the consulting services of a technically skilled programmer or administrator.

Freemium strategies were common among the most popular open-source programs; they thrived when many customers wanted additional services. For example, Red Hat, founded in 1993, installed the Linux operating system on computers and offered consulting services for businesses seeking help creating a computer suite based on open-source programs. IBM acquired Red Hat in June 2019. Nginx, the most common Webserver for streaming on the Internet, offered similar options. The basic Nginx webserver was available for free, but large-scale deployments worked more reliably with additional services, and many users desired these features.¹⁵

In 2019, the revenue of open-source-based services was estimated to be \$17 billion, and by 2022, it was expected to grow to \$30 billion.¹⁶

Origins

In 2016, Delangue, Chaumond, and Wolf were ready to move on from their prior startups and try something new. They were already friends who had discussed starting a business together and were excited by ML's promise for offering new ways of building technology.¹⁷

Fun, Conversational AI?

In a field with many startups, the three considered various projects, eventually settling on developing a conversational AI chatbot. Wolf recounted, "Chatbots were all very hyped and everyone was doing them, yet they were built with very, very, simple technology. We set out to see what recent advances in AI could be used in chatbots." Delangue added, "At the time, Apple's Siri and Amazon's Alexa were really boring, very productivity-driven, not really conversational. We thought, OK, let's build a fun, conversational AI." They named the company Hugging Face after the hugging face emoji, one of many stylized emoji available for users to send as part of text-based communications. It signaled the friendliness they sought to develop in their technology (see **Figure A**).

Figure A Hugging Face Logo



Source: Company documents.

To create an effective conversational program, the Hugging Face team curated a broad range of data sets and models to solve conversational problems across diverse topics. The chatbot technology exceeded their expectations, winning the second Conversational Intelligence Challenge (ConvAI2) at NeurIPS 2018, a well-known ML conference.

Motivated by a goal of creating visibility for Hugging Face and by an internal sense of mission, the team also published their work in scientific conferences and made their code available under an open-source license. "Our motivation comes from different places," said Wolf. "For me, it came from my

physics background, where it was really hard to get access in physic research, and I wanted to fight against that. For Julien, he's from an open-source background and felt it was an important thing." Their best-known library, Transformers, made ML models accessible, and was released in April 2018 under the Apache License 2.0,^b which allowed collaboration in any software ecosystem.

As a happy accident, their decision to open-source their technology prompted a community to begin forming around their software. Transformers was valuable because it adapted recent research and made it easier.^{c18} Delangue said, "We were releasing things without thinking too much about it, more as a way to pay it forward for the community. But, very quickly, the community started to gather around the platform, started to contribute like open-source contributors, and then progressively they started to share their models."

Initial Funding

The founders' initial business plan was to generate revenue from their next-generation chat bot. They were unconcerned with an immediate monetization plan for the bot, aiming instead to create a wide user base, as social media platforms often did. In 2018, a reviewer wrote, "[T]he startup has been building a chatbot app with a strong personality for bored teenagers. Instead of focusing on customer support or convenience, Hugging Face is focusing on emotions and entertainment."¹⁹

The team looked for investors both willing to invest in a company without an initial monetization plan and willing to give them the latitude to experiment and innovate. Wolf recalled, "Not every VC was interested or would have been great as a base investor. We said 'no' to many people who wanted a short-term monetization plan—it would have just constrained us too much in terms of scope." In March 2017, Hugging Face raised an angel round of \$1.2 million for their vision to develop their chatbot.²⁰ However, by 2019, they faced a choice about the direction of their company.

Pivot to Open Source

Although their chatbot technology was effective, they were surprised at the amount of growth they were observing in the use of their open-source technologies. Said Wolf, "We saw the exponential that you always want in your startup—exponential in the number of users, exponential in the number of installs of the library. All this traction around open source aligned with some deep thing we believed in. And so we decided to pivot."

The team rallied around a new end-goal of creating a coherent open-source ML ecosystem that nurtured others. The decision to develop around an open-source community changed how the team approached development. Rather than hiring engineers to work on ML models directly, Hugging Face assembled a team to provide a framework within which the broader community could work. They created a website to make it easier to find and access the technologies that were available through their libraries. To complement Transformers, the team added a second open-source library, called Datasets, which aggregated data used to train models—making it easier for others to train models. The original chatbot became one of many projects available at Hugging Face. Said Delangue, "The underlying abstraction that we built for the chatbot became the early innings of what Hugging Face is today."

^b Similar to many open source licenses, the Apache 2.0 license permitted developers to change, copy, or update source code. It also required the developer to make any copies or modifications available to others.

^c As of May 2022, the Transformers library had 62,000 stars and 14,000 forks on GitHub. GitHub stars and forks were general measures of project popularity. Each star represented a user that bookmarked the repository for future reference. A fork occurred when user copied the source code of the library into a separate repository, either for inspection or for custom development.

Hugging Face's community-first approach to development contrasted with that of larger technology companies. Wolf explained:

We try very hard not to reproduce the mistakes of some open-source ecosystems in ML. For example, Google just makes software for itself. It's hard to care about people outside of Google when you're so big. Usually, these new large AI models make it difficult to actually do research on this model, and try to understand how they work, where they don't work, how to improve them. They are only studied by the team who created them, which is like five or six people. So it's a very small foundation.

Growth

Hugging Face's platform continued to grow exponentially, with the company continuing to add capabilities toward creating an end-to-end open-source ML ecosystem. From an employee count of 38 in 2020, by the late spring of 2022, Hugging Face had grown to more than 150.

The growth was evident on the website. Hugging Face operated a website with 30,000 datasets, 100,000 models, and 20,000 Spaces (or demos). Its most common model solved problems in text classification, but other models supported question answering, image classification, translation, summarization, and object detection.^d (See **Exhibit 4** for company landing page.)

An Open-Source Machine Learning Company

In spring 2022, Hugging Face organized its team around three core priorities — developing the main platform, supporting the open-source community, and pursuing cutting-edge scientific research. The team's vision for a coherent open-source ML ecosystem was realized as a multi-sided platform, connecting users (i.e., researchers, companies, and individuals) to range of ML technology resources.

The Hugging Face Platform

The resources on the platform varied in their distance to practical application, and ranged from datasets and untrained model architectures to fully functional model demos. Its most popular technologies were pre-trained models — models that could already solve general-purpose ML problems (such as parsing language or images), but could be adapted to specific use-cases via fine-tuning — a simpler method than training a model from scratch. These models could be applied to general NLP use-cases (such as extracting the emotional sentiment of a sentence), or to extract word embeddings for other models to solve user-specific use-cases. For example, a model originally designed to process text-based inputs and respond to a prompt could be modified into a customer service bot that processed queries related to a company's products. This could help a firm direct its customers to the answers they were looking for in a more efficient and cost-effective way.

Hugging Face also hosted an extensive collection of datasets that could be used to train or expand ML models; most of these were sets of text-based data and were referred to as Datasets. The data sets were diverse: one had strings of text tagged as either grammatically correct or incorrect, which could

^d Text or image classification assigned a label or class to a given text. Question and answer models retrieved answers from a given sentence, usually in question format. Translation converted text from one language to another. Summarization produced a shorter text describing the information of a longer document. Object detection identified objects within visual images. Source: Hugging Face, <https://huggingface.co/>, accessed October 2022.

be used to test models for their grammatical accuracy. Contributors imported many of these datasets out of a desire to share resources with Hugging Face's open-source community.

Hugging Face's resources could be accessed in two ways: via the user-friendly website, and the more technical software libraries (including Transformers and Datasets). Hugging Face's libraries were extensively documented, making it easy for new users to get started using their technology. The website offered users guided access to the libraries and to the models and datasets created by the open-source community. While most of the model code was hosted on GitHub, the model weights^e lived on Hugging Face's website, making it easier for users to access and train model weights. Furthermore, the website was extremely usable and focused on ML, making it the go-to location to search for and identify new models. The website also offered Spaces where users could interact with demos directly through their Internet browser. Each of these components was created to inspire users to create and share their own work.

Delangue explained, "There's an easy flow between these three components; they inform and build off one another. For instance, someone can visit the site and see a popular Spaces demo, discover which models were used, pick a new dataset to fine-tune it with, generate a new model, and (maybe) build a new Spaces demo." Technical users could access the same resource through Hugging Face's software libraries—packages of related resources and functionalities organized by topic. To Delangue, the datasets, models, and Spaces informed and reinforced one another (see **Exhibit 5** for illustration).

Use of the platform remained free and unrestricted—any user could access any model and data set and download, use, and modify it. Delangue saw advantages in this approach. He said, "With open source, you don't have to capture 100% of the value you create. The platform creates 1000x more value being open than if it were closed. Even capturing 1% of this value has the potential to be more successful than the traditional proprietary approach."

Guiding an Open-Source Community

Even in 2020, when the Hugging Face team had just over 30 employees, their libraries and models attracted a group of dedicated volunteer contributors. At the end of 2018, 10 teams contributed monthly to Transformers; by 2020, that number had grown to over 30 monthly team contributors. Beyond contributors who directly contributed to the library codebases, many individuals and organizations contributed new models to Hugging Face's model hub for any other user to download and use (see **Exhibit 6**). The contributors were distinct from the more significant number of users who used the Hugging Face models but did not contribute new code or models. The company managed through Slack and discussion forums; most meetings were conducted asynchronously to promote inclusivity.

Contributors themselves were often experts, and open-source culture encouraged them to share that knowledge with other academics, professional computer programmers at other companies, or independent coders. This enabled Hugging Face to keep up with NLP and ML technology, which evolved so quickly that it was often impossible for a company to keep up. Hugging Face Open Source lead Lysandre Debut, said, "If we have ten experts in machine learning, the field moves so fast that our experts cannot implement everything that comes out. Enabling our community means that we can keep up to date, because an interested community member can work on new technology, provide it as a tool, and make it make sense to the rest of our community."

^e Model weights are numbers that encode the patterns models learned from training data. The same model might have different weights if it was trained on different datasets.

In a typical arrangement, Hugging Face employees organized and supported the code and data that contributors produced, offering suggestions and creating tools. This led Hugging Face to expand into supporting other areas of ML, beyond NLP. In February 2021, Hugging Face added audio processing features, and in March 2021, the firm added vision processing. Debut said, “The community really appreciated us expanding the field into audio and vision – it didn’t translate into usage right away, but a lot of people talked and tweeted about it.”

The company aimed to prioritize their contributors’ involvement over those of Hugging Face’s employees. If an employee and contributor created similar models, the contributor’s model usually won out in order to encourage volunteer contributions. Debut said, “We’re always devoted to helping that community contributor achieve their work, and we hope that in doing that, they learn and are happy, and will come again to contribute once more.”

Along with community proposals, every six months Hugging Face surveyed its contributors to gauge satisfaction and interest in new features. This helped the leadership team validate company plans. A community survey from April 2021 showed that Hugging Face drew substantial usage from NLP researchers, data scientists, and ML engineers (see **Exhibit 7**). Most users liked the Hugging Face libraries and recommended them, with more satisfaction for the integration with PyTorch compared to integration with the TensorFlow.²¹

Hugging Face also organized sprints^f to focus short-term employee and contributor experiments. For example, in 2021, Hugging Face employees coordinated a sprint focused on using JAX, an open-source numerical computing framework created by Google researchers. The sprint involved groups of five to 10 employees and volunteers, similar to a hackathon, and included 800 participants and many projects over one month.

One project to come out of the JAX sprint was DALL-E mini, a test program that sought to replicate an earlier program from ML company Craiyon. DALL-E mini created images that tried to match text-based prompts. (See **Exhibit 8** for sample prompt and images.) The DALL-E mini program interested those who worked on it, leading Hugging Face to continue the experiment for over a year. In spring 2022, the DALL-E mini team released the program to the public, where it gained an enormous number of users in a short time. A couple of months later, Delangue recalled:

Someone from our community kept working on the model, uploaded it on the platform, and it went crazy viral. I think it generated like 200 million images in the past seven days. That’s something that we wouldn’t have prioritized as a team and that we might have missed, but the community enthusiasm meant that someone came up with it anyway. This enables us to adapt to the changing topics and to stay at the cutting edge.

Adding Enterprise Features

Hugging Face approached monetization opportunities with caution, concerned about the tension between maintaining the open-source ethos of its community and charging to support platform development. Delangue said, “That strategy definitely served us well, and looking back I’m really happy we did it that way rather than focusing too early on monetization.”

Charging for basic use of the platform was out of the question, as it would deter the volunteer contributors who made Hugging Face so popular. However, Hugging Face received many requests for

^f Sprints were periods of time for designated work tasks, a strategy taken from Agile management systems. Sprints usually ran several weeks, with the assumption that this helped people focus on quickly solving specific problems.

advice from well-known industry experts from major companies that were starting to incorporate ML into their products. Companies wanted help incorporating Hugging Face's features in their own infrastructure for increased security and customization. In 2021, the company began to charge for access to some of its experts and for special explanations of new features.

These features focused on accelerating AI deployment and usage by simplifying the MLOps management for companies. Hosting a production ML model posed many challenges beyond selecting a specific model type, and managing the model lifecycle in an ideal AI factory (refer to **Exhibit 3**) was a significant challenge for businesses lacking strong ML skills. Hugging Face responded by creating functionality and interfaces that would store, train, and apply models to custom use-cases. They also offered consulting services for companies that wanted to use AI for their data but did not know how or where to start. Plans started as low as \$9-per-month for limited usage, but pricing was adjusted to match enterprise size and needs. Pay-as-you-go and custom options were available for companies looking for more hands-on assistance or to incorporate larger amounts of data. (See **Exhibit 9**.)

By 2022, Hugging Face had more than 1,000 paying clients, mostly comprising companies that sought to employ AI and NLP tools in their products or operations but lacked the capacity to create or manage the technology in-house. A large majority of the clients arrived organically via inbound requests. Many were familiar with Hugging Face by name or from the company's active ML community. They typically approached with the intention of learning more about Hugging Face's services, which led to discovery calls and eventual sales. Jeff Boudier, Hugging Face's head of product and growth, said, "I start any sales conversation by understanding what the customer is trying to do, and what would be the best way for us to make them more successful. Even organizations that have data science teams are just using the tip of the spear compared to people who are experienced in applying this. So we have a lot of value to offer, to unlock and make sure that they get to success."

The most common enterprise applications for clients focused on classification tasks based on text, audio clip, or video clips. Chaumond noted, "The majority of use-cases in industry are classification problems – document classifications or token classification to do information extraction. We found that the most efficient way to solve a classification problem is to fine-tune a classification model."²² However, most clients used many models, making for long, continuous relationships with Hugging Face. In 2022, Hugging Face's paying customers included Bloomberg, Intel, Grammarly, and Pfizer.

Company Operations

Hugging Face's operating model emphasized employee ownership and initiative.

A Company of Generalists

From the start, Hugging Face's team wanted the company to grow naturally, without rigid structures or job titles. Delangue recalled, "We started with the idea that people are much more multifaceted than large organizations allow them to be. We try to turn that into a strength for the company, saying to people, 'It doesn't matter that your main job is to be a developer, researcher, finance person, a salesperson. You can also do communication marketing, or research, or anything.'" This generalist mindset dramatically impacted the way the team structured roles across the organization.

Company structure Hugging Face was a remote organization. In 2022, the three founders were based in Paris, Miami and the Netherlands. Most employees worked remotely, around the globe. The company supported a few small offices in cities where employees were colocated, including Bern, Montreal, New York City, Paris, and San Francisco. Because employees lived in many different time

zones, it was difficult to schedule large meetings. Instead, teams relied on Slack for internal communications. For announcements and conversations with contributors and users, the company relied on Discord, GitHub, Hugging Face's community tab, online forums, Slack, and social media.

Rather than organizing around traditional functions, Hugging Face organized around three focal points: revenue, visibility, and usage, each of which had a different team working on its respective focus without regard for the other two (see **Exhibit 10**). Debut explained: "The business team is focused on revenue. The science team is focused on visibility, and the open-source team is focused on usage. In the open-source team, we're focused on usage rather than revenue. This separation of concerns enables our open-source work to be truly free and open and does not reward putting important features behind paywalls. We, of course, collaborate with other teams, so we understand our revenue streams and business perspectives, but at the end of the day our focus is on raw usage."

Shared responsibilities Within each area of focus, the organization was flat, with a few designated leaders. Employee skillsets varied but all were encouraged to adopt a generalist mindset with regard to the tasks that they undertook. For example, employees were responsible for their own marketing. The company encouraged informal communication to bring attention to their activities. Delangue said, "We don't have anyone specifically focusing on marketing full time, so we tell even our most specialized researcher or developer, 'Marketing is your responsibility as much as everything else. When you build a new product, a new feature, a new technology, you should spend time tweeting about it. You should spend time talking to journalists about it.'"

Similarly, no single community monitor was responsible for responding to inbound messages from users or other companies. Omar Sanseviero, Hugging Face's ML engineer lead, said, "That means that people from other teams or from the development workers' engineering team can really jump into any discussion and work with them. But there is no single person in charge of it. It's really like a distributed, shared responsibility."

Hugging Face diverged from its generalist structure in only a few areas. It did not maintain its own computer servers, relying instead on an internal team to contract with and manage the use of cloud computing services for the space they needed. Regulation meant that some logistical functions required more focused expertise: "Finance and legal were the first functions that we needed infrastructure for, so we have one full-time finance person and one full-time legal person," said Delangue.

Hiring Its broad base of contributors meant that Hugging Face did not have to hire experts in every aspect of NLP and ML. However, if the founders noted high interest in a set of models or a new library, they looked to bring an expert in that area into the company.

Hugging Face had developed a reputation for being an exciting and prestigious place for open-source developers to work, and the team found it easy to attract new employees. Potential employees had to be comfortable working for a company where almost all work was self-determined and almost all conversations were public, including discussions of project failure. Hiring was also decentralized. Debut said, "If a team member has someone that they really want to work with, then we're super happy to hire them. If there is someone who really seems like a good hire, then we're never going to pass on that person. So even if we're not actively hiring, I'm still checking all candidates basically on a daily basis." No employee filled just one role; Boudier added, "Hiring people who are willing to adapt to whatever needs doing makes hiring dynamics more fluid and reactive."

While some contributors aspired to become employees, Hugging Face discouraged the idea that volunteering as a contributor could be seen as an audition process for a job; most contributors remained volunteers. Still, Hugging Face often hired contributors to hire for new positions, since the team was

already familiar with their work and knew that the contributor understood the company. The transition to becoming an employee required a shift in perspective. Debut explained, “As members, someone would contribute models or model checkpoints. As employees, they would be creating ways to contribute models and model checkpoints. So, they’re going a step above in the model stack to make it easier to perform machine-learning engineering without actually performing a lot of the machine-learning engineering.”

Proposing New Projects

Anyone in the Hugging Face community – either employee or contributor – could propose a new project. Projects could take the form of research projects or new features for the Hugging Face platform. The typical project involved a team of employees and contributors, although sometimes incubator projects had a single contributor. Management had no formal mechanism to shut down a project.

However, while employees decided what projects to pursue, management buy-in was key to gaining internal recognition for the project and attracting the resource support of other community and internal members. Buy-in could be gained through a project proposal. Depending on the project, the team, and the context, proposals might take the form of a Google document, Notion post, Slack, or other asynchronous communications. At times, the process pivoted to synchronous meetings if it seemed the medium created more work than benefits. Wolf elaborated, “Often people were happy to continue to iterate on the Google doc, instead of working on the actual idea. You can make almost anything sound interesting in a proposal, and once you’ve written it, you lose the flexibility of the project itself. We changed it so that in those instances we talked to people face-to-face.”

The management team viewed themselves as advisors and connectors rather than having formal power to approve or shut down projects. Instead, they provided honest feedback on whether they thought the project worthwhile or ambitious enough. Wolf said, “People come to me ten times a day with a proposal for a new project, and I say yes to maybe 1% of it. We also do a lot of small bets with maybe one or two people, to see if they can prove that it will work. We want people to have a plan – what happens if they succeed? What happens if they fail? What are the lessons?”

Delangue knew there was a delicate balance to be maintained between leading and keeping the company decentralized. He said:

We want to let other people lead, but to keep everyone aligned. One of the problems that we run into is that if an organization is super decentralized, people are going to go in all directions because no one is making decisions. But with feedback and communication, someone is able to look at a new proposal and say, “I’ve tried that five times in the past. I wouldn’t waste your time with it, and here’s why.” That creates a system that so far has worked really well for us. We have radical transparency and knowledge sharing, while still having the involvement of people who feel ownership over their products.

The team viewed this model as empowering employees to do their best work. Wolf said, “When people are driven, when they’re excited about their project, they don’t really need you – they’re happier when you’re not there.” Debut agreed, saying, “The most important pillar is always community – we’re educators and helpers of the community, less than instigators.”

Once a project was chosen, the team and the community completed work much faster than traditional organizations did. Delangue said, “The ability to correct course as a decentralized organization is usually very short. We might miss something for a few weeks, but then we’ll see a problem very quickly. The time between deciding to create a new project and having a testable version

is very short.” Although the company was still small, the pool of contributors meant that much of the work was done quickly. Sanseviero said, “One superpower of the company structure is that each employee can have the impact of five or 10 people because they can leverage a team of volunteers.”

The managers considered failure an important aspect of understanding what projects would work, and employees who had previously led projects that failed were more often chosen to lead future projects. Wolf said, “When an employee has failed at a project, we feel like you can trust them more, because they’re better at identifying products that fail or work well – they get an intuition that someone where everything was successful often doesn’t have.”

Continued Funding

Hugging Face’s business model attracted many investors who were interested in a new, rapidly-growing ML platform. GitHub’s success, and its purchase by Microsoft in 2018, helped Hugging Face frame its financing conversation. Wolf said, “We raised money around this idea that, while Hugging Face was not exactly the GitHub of machine learning, open source-powered machine learning could be a big thing in the future and attract a similar type of community.”

Angel seed funding in March 2017 and May 2018 was followed by a Series A round in December 2019, a Series B in March 2021, and a Series C in May 2022 (see **Exhibit 11**). Boudier explained:

We’re very forward about our values and what we’re trying to do. There’s no secret – investors that put money into Hugging Face are bought in and are necessarily long-term thinkers. We have a shot at building a category-defining company in a new technology, and we can only do so if we’re able to experiment and think long-term, which means we need money in the bank for 12 or 18 months from now. We don’t want to suddenly have to reevaluate how we do things because we’re running out of money. Having Series C money gives us the confidence that we have time to realize our opportunity.

Hugging Face in 2022

By spring 2022, Hugging Face had grown to 150 employees. Delangue was happy with the company’s size as it stood, saying, “This is the extent that I can see in terms of scale. The way we operate is the same as when we were 40 or 50 people – it might work if we were 250, but not 1,000.” About 30% of team members focused on science efforts, 40% on open-source and developer advocacy, and 30% on the business. There were 1,000 frequent contributors, 11,000 people on Hugging Face’s Discord server, and, over 100,000 daily users.

The company continued to pursue priorities across its three focal areas: platform development, supporting the open-source community, and pursuing cutting-edge scientific research (see **Exhibit 12**). For the platform itself, in addition to adding enterprise features, the company was interested in adding open-source implementations of some recent research papers focused on tabular data. For its open-source community, the focus was on expanding support for ML models beyond NLP. In its research efforts, the company contributed to the area of Large Language Models (LLM) by releasing one of the first purely open-source LLMs – BLOOM.

A deep learning approach to tabular data had perhaps the most intriguing commercial value because many companies typically stored data in this form. Tabular data referred to data stored in rows and columns, like a spreadsheet, rather than as the images or text that formed the focus of most frontier ML research. For Delangue and his team, tabular data provided new opportunities, but also some

potential drawbacks: it was not a core product for Hugging Face, and the community found other ML topics more exciting. Yet tabular data analysis had the potential to attract many more paid users for Hugging Face. Sanseviero said, “We know that everyone is using tabular models, but not at the same rate that they’re using other types of models. We currently have two team members who are creating some nice open-source tooling to allow anyone to share their models with others. But it’s an early stage for us as a developer advocacy group in that we don’t have huge adoption by the community.”

Delangue was keen to continue the current operating model. He said, “What I want is to get to a place where I don’t have to tell anyone what to do, where people can decide for themselves what to work on, and I just formalize and spread the culture.” Yet, he was interested in finding new ways to monetize the company’s capabilities, especially when some—including tabular data—could dramatically increase the company’s earnings. Wolf agreed, “We still haven’t nailed the part that will be the most interesting in terms of value.”

As NLP and ML technology continued to develop, Delangue and his team faced several options of what to focus on next. Wolf said, “Because our company is so adaptable, we could do almost anything in AI, but we could also go outside of AI. We could do some interesting machine learning for climate change, for molecule discovery, for social science—anything where you use a computer. The scope is gigantic. The main difficulty is getting people to think big.”

An Unexpected Contribution

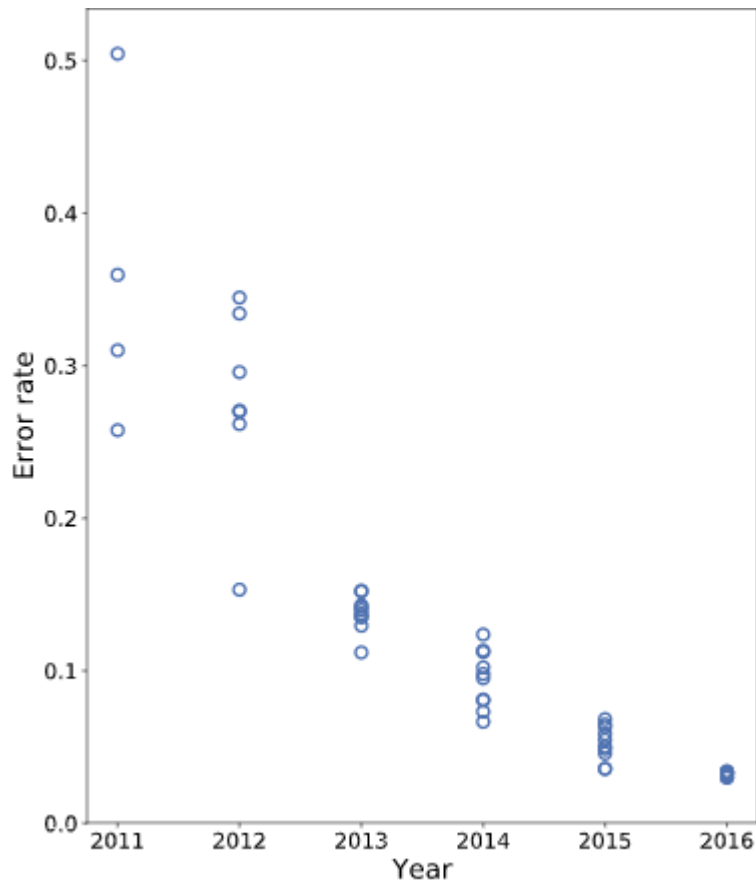
In May 2022, a community member uploaded a model called GPT-4chan to the Hugging Face website. GPT-4chan took NLP research in a controversial direction. It was an advanced language model trained on the website 4chan, an online discussion board with a reputation for hosting illegal and offensive content. Posts from a discussion organized under “Politically Incorrect Board” had trained the model over three and a half years.²³ The results mimicked the site’s speech, displaying a strong tendency to produce toxic, racist, and derogatory language. The contributor explained:

GPT-4chan is trained on anonymously posted and sparsely moderated discussions of political topics. Its intended use is to reproduce text according to the distribution of its input data. It may also be a useful tool to investigate discourse in such anonymous online communities. Lastly, it has potential applications in tasks such as toxicity detection, as initial experiments show promising zero-shot results^g when comparing a string’s likelihood under GPT-4chan to its likelihood under GPT-J 6B.²⁴

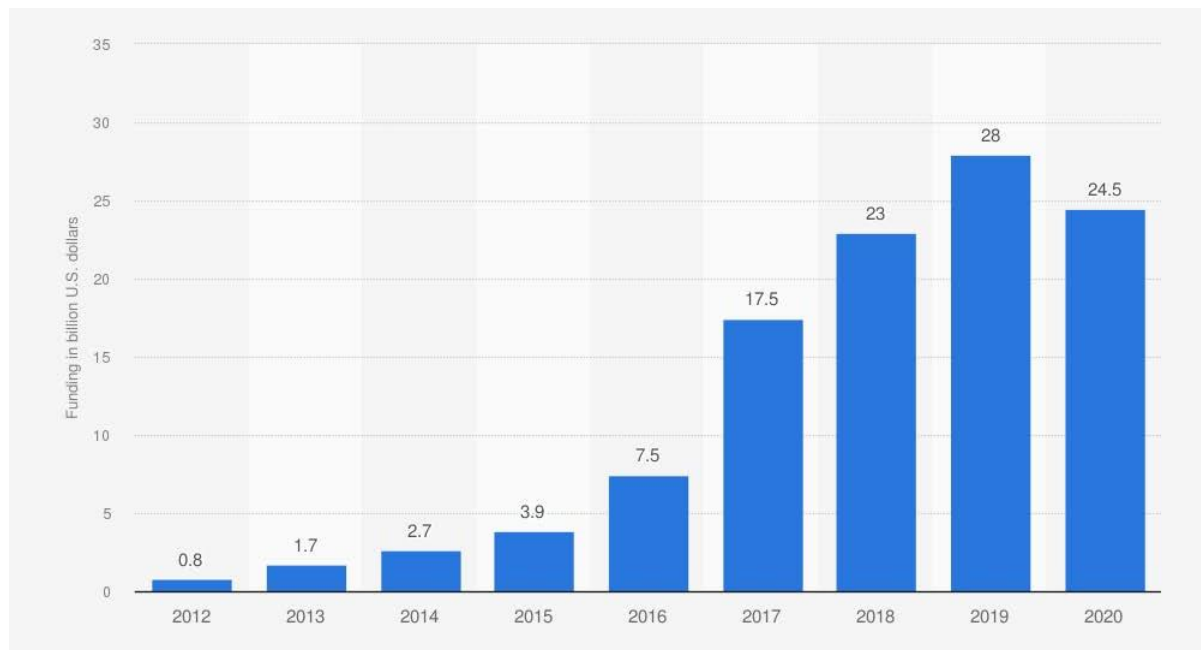
After some discussion, Hugging Face disabled access to the model on its website, but kept the debate and documentation available. Looking back, Delangue recalled that many internal opinions had defined sharply contrasting points of view.^h On the one hand, allowing GPT-4chan on the platform would uphold the ethos of openness that enabled civil discussion and research that Hugging Face had worked hard to create. Yet the model had already generated harmful content. The majority of the Hugging Face community had called out the limited research value and clear dangers of continuing to support the model. Blocking the model would protect the community’s values and prevent Hugging Face from being part of potential harm. In fall 2022, with the benefit of a few months of hindsight and a calmer moment for reflection, had Hugging Face made the right decision?

^g Zero-shot results were a setting in which a program yielded results without training data. In context, the author was claiming that GPT-4chan generated better results than a comparable alternative for detecting toxic language.

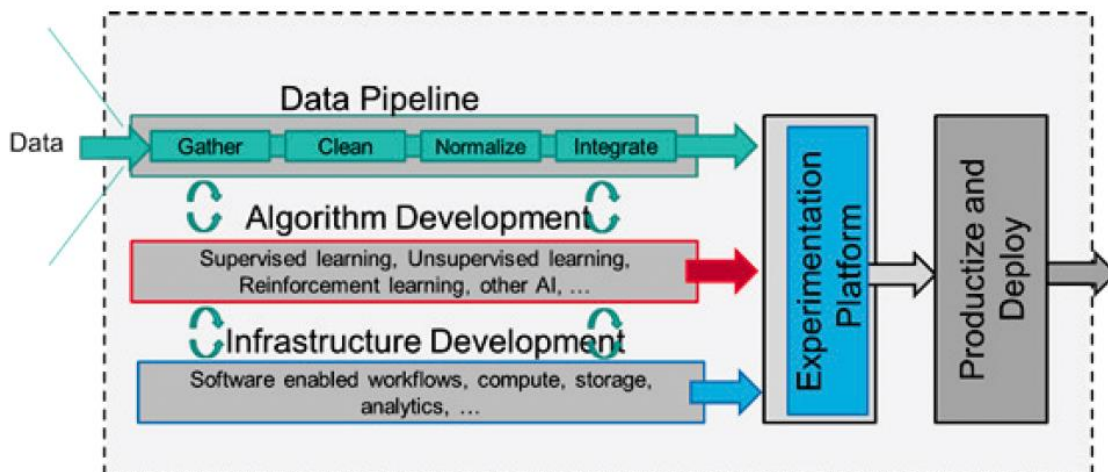
^h For contrasting points of view, see <https://huggingface.co/ykilcher/gpt-4chan/discussions>, accessed October 2022.

Exhibit 1 ImageNet Large Scale Visual Recognition Challenge Results, 2010-2017

Source: By Gkrusze, Own work, <https://commons.wikimedia.org/w/index.php?curid=69750373>, [https://commons.wikimedia.org/wiki/File:ImageNet_error_rate_history_\(just_systems\).svg](https://commons.wikimedia.org/wiki/File:ImageNet_error_rate_history_(just_systems).svg), accessed January 2023. Licensed under [Creative Commons Attribution-ShareAlike 4.0 International \(CC BY-SA 4.0\)](https://creativecommons.org/licenses/by-sa/4.0/).

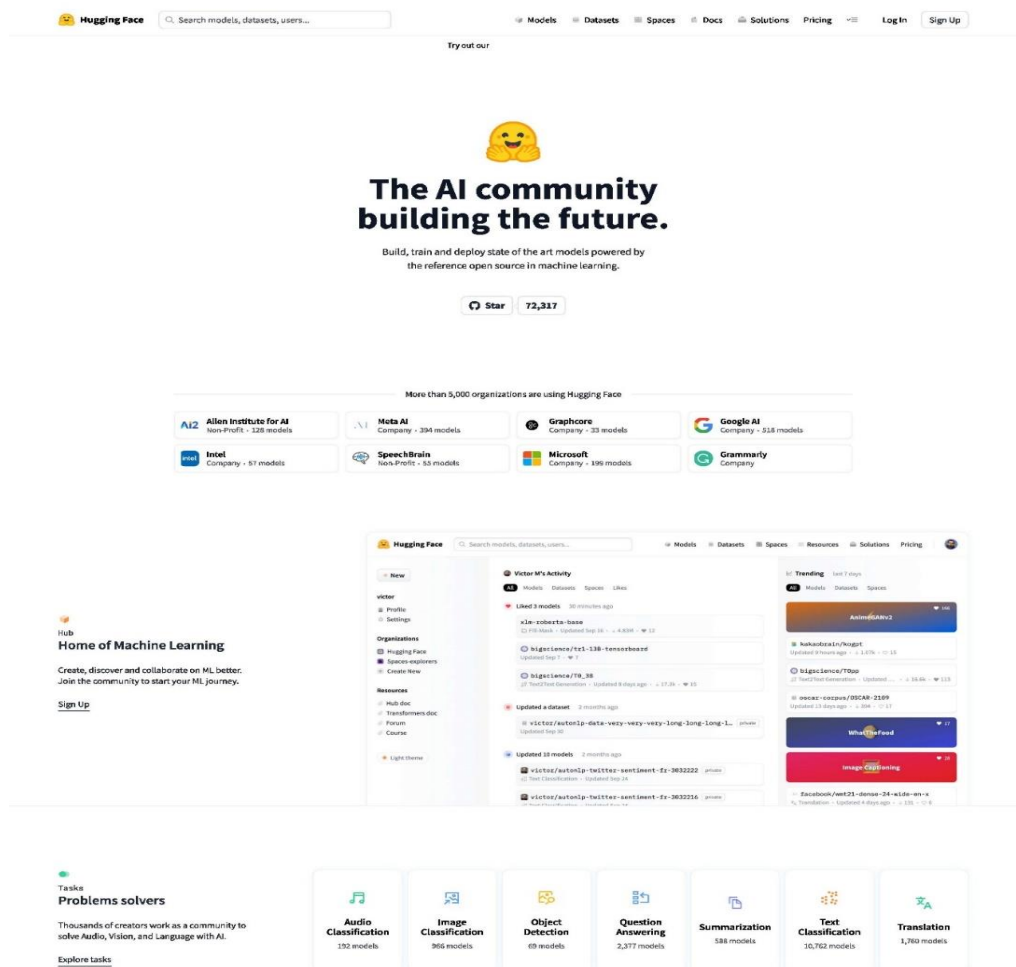
Exhibit 2 AI Startup Funding Worldwide, 2012 to 2020 (in US\$ billions)

Source: Berger Thormundsson, "AI startup funding worldwide, 2012-2020," Statista, March 17, 2022, via Statista.

Exhibit 3 Sample AI Factory Diagram

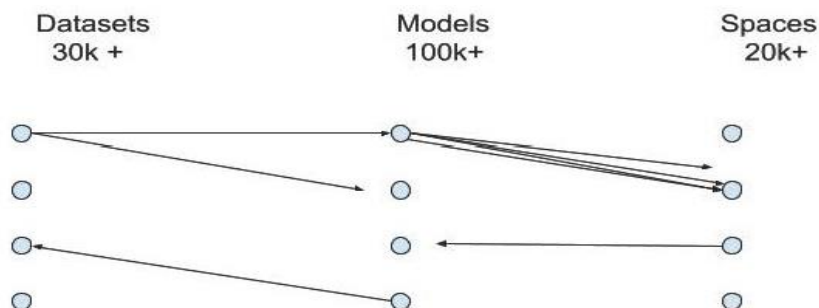
Source: Marco Iansiti and Karim R. Lakhani, *Competing in the Age of AI: Strategy and Leadership When Algorithms and Networks Run the World*, (Boston: Harvard Business Review Press, 2020), p. 53.

Exhibit 4 Hugging Face Home Page

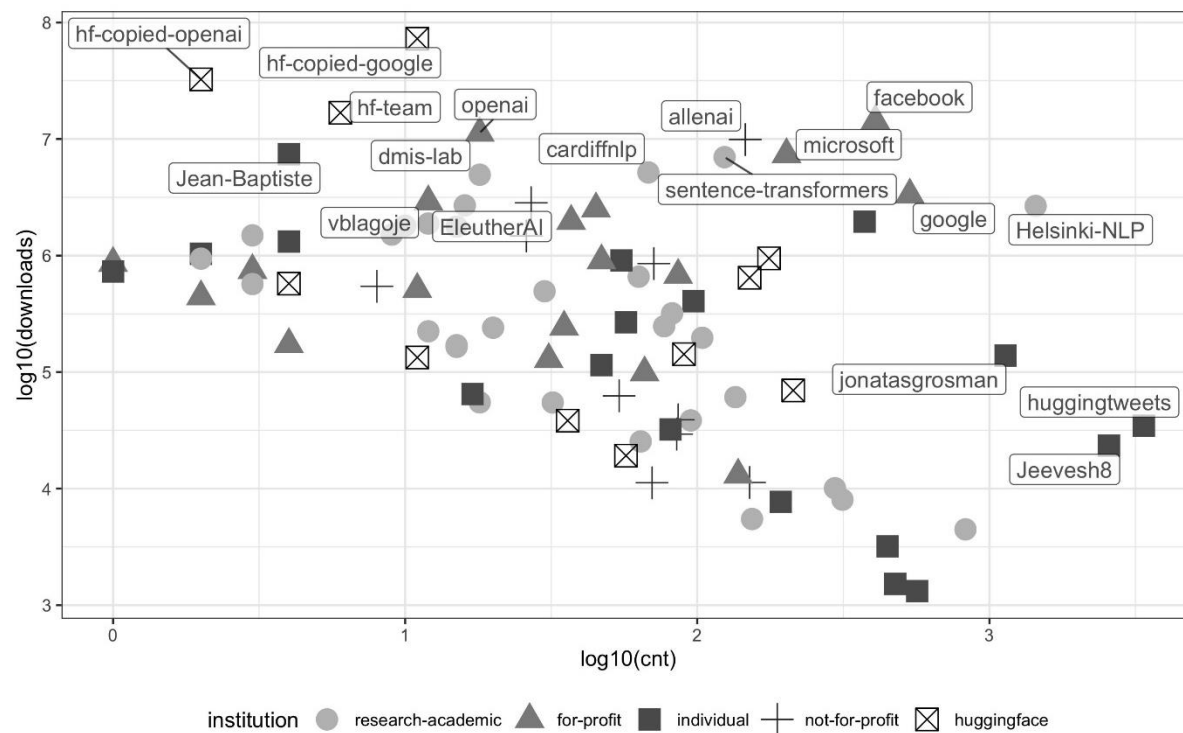


Source: Company documents.

Exhibit 5 Illustration of Hugging Face Resources



Source: Company documents.

Exhibit 6 User Downloads (Y) by Models Uploaded (X), September-October 2022.

Source: Casewriters, based on public usage data from Hugging Face.

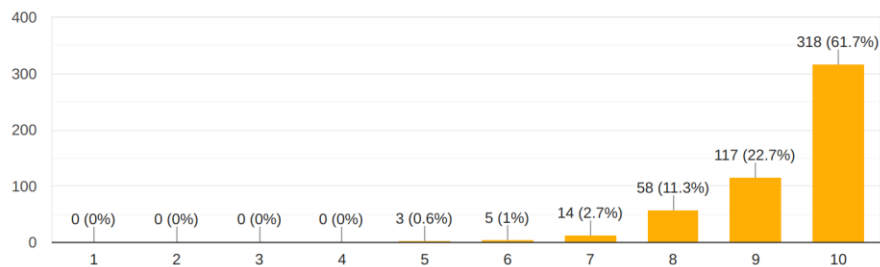
This plot visualizes Hugging Face's contributors, based on number of models uploaded (x-axis) and number of times those models were downloaded and used by other users at Hugging Face (y-axis), based on the type of contributor (greyscale).

Sample: While, as of November 2022, there were more than 24,000 contributors on Hugging Face, here, we show only the top 75 contributors (as measured by having more model uploads or user downloads). This sample was selected to show the most influential or intense contributors, and illustrates the significant variation in the type of contributions made by the Hugging Face community. Therefore, it is intentionally not representative of the typical contributor on Hugging Face.

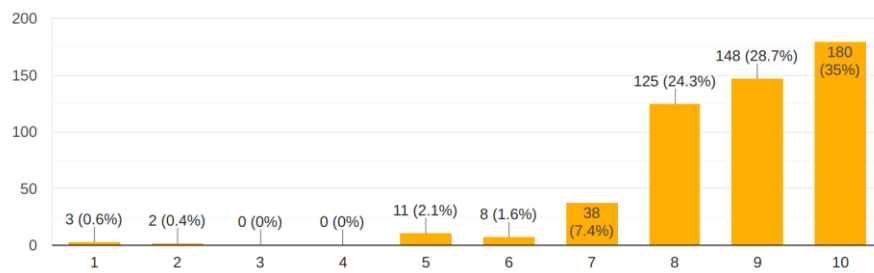
Some notes on how to interpret the data: An uploaded model was one that was listed on the Hugging Face model directory (<https://huggingface.co/models>). Any time a user called any of those models on the Hugging Face hub using the Transformers library it recorded a model download. Here, we show count of model downloads from the period of September 27, 2022 through October 26, 2022. The institutions labelled were based on the listed contributors on the Hugging Face platform (e.g., <https://huggingface.co/allenai>), except for the 'huggingface' institution (which was manually coded by the case authors).

Exhibit 7 Community Survey Results, April 2021

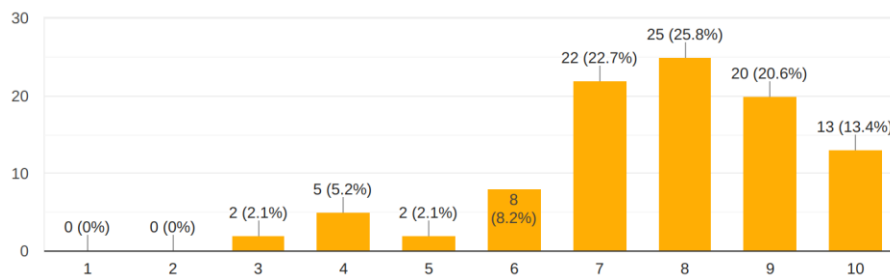
Recommending the library



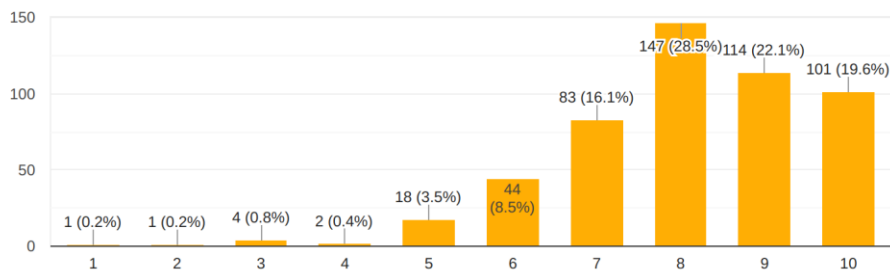
Recommending PyTorch Side of the Library



Recommending TensorFlow Side of the Library

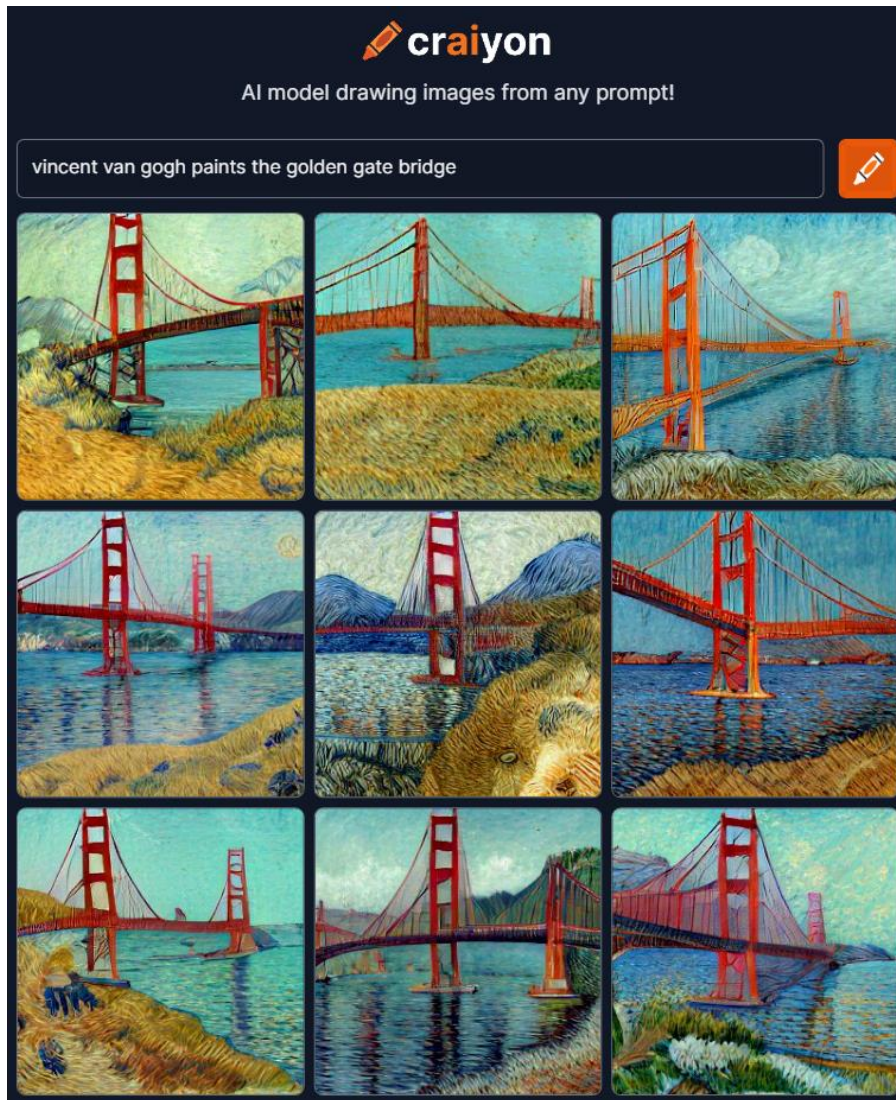


General sentiment regarding documentation



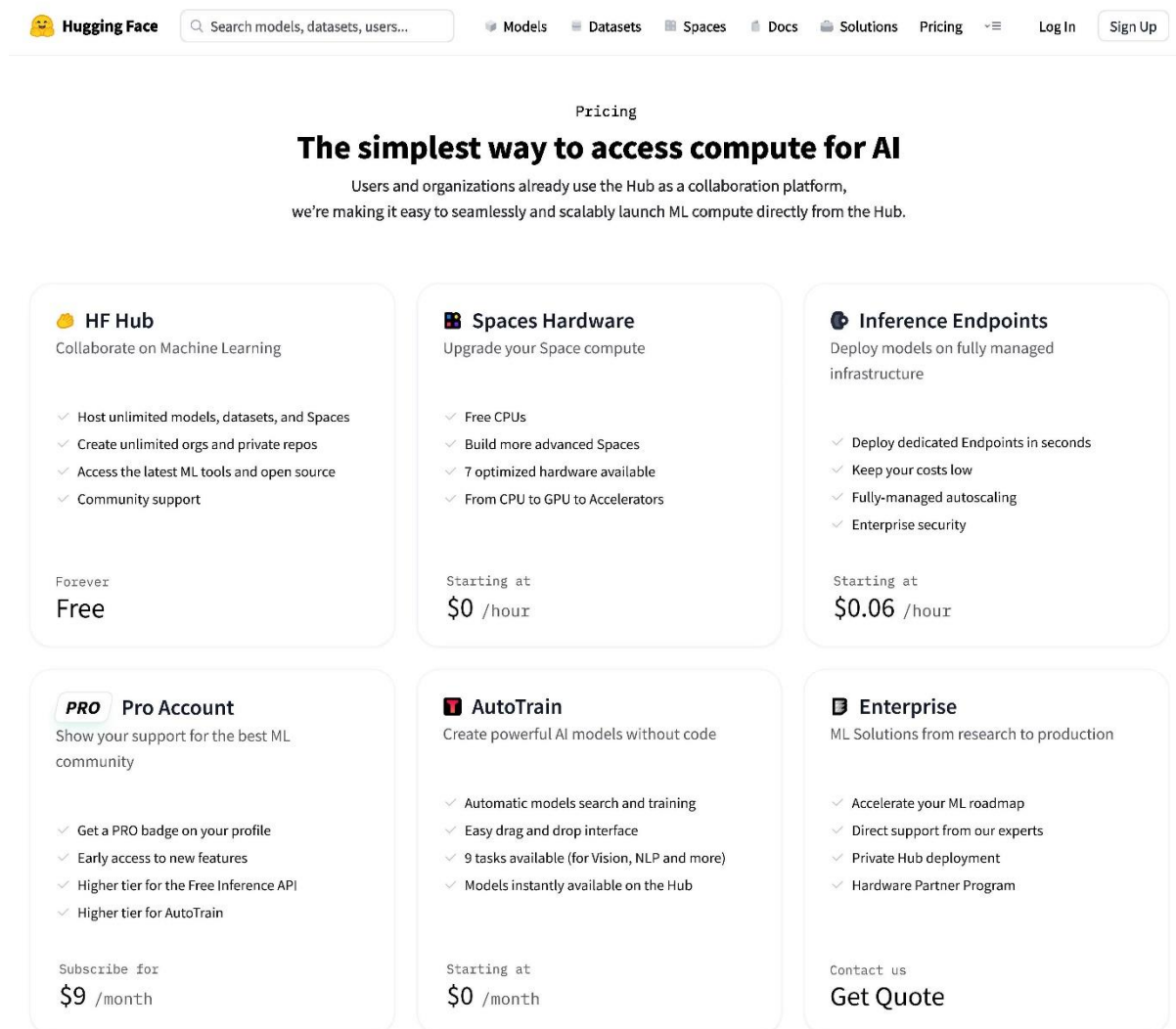
Source: Hugging Face, “Transformers Huge Community feedback: 40k,” April 2021, <https://discuss.huggingface.co/t/transformers-huge-community-feedback-40k/5313>, accessed October 2022.

Note: Survey participants were asked to rate each question on a scale of 1 to 10, with 10 indicating highest level of favor.

Exhibit 8 Craiyon.com Example

Source: Craiyon.com, accessed October 2022.

Exhibit 9 HuggingFace: Pricing Home Page



Hugging Face Search models, datasets, users...

Models Datasets Spaces Docs Solutions Pricing Log In Sign Up

Pricing

The simplest way to access compute for AI

Users and organizations already use the Hub as a collaboration platform, we're making it easy to seamlessly and scalably launch ML compute directly from the Hub.


HF Hub
Collaborate on Machine Learning

- ✓ Host unlimited models, datasets, and Spaces
- ✓ Create unlimited orgs and private repos
- ✓ Access the latest ML tools and open source
- ✓ Community support

Forever
Free


Spaces Hardware
Upgrade your Space compute

- ✓ Free CPUs
- ✓ Build more advanced Spaces
- ✓ 7 optimized hardware available
- ✓ From CPU to GPU to Accelerators

Starting at
\$0 /hour


Inference Endpoints
Deploy models on fully managed infrastructure

- ✓ Deploy dedicated Endpoints in seconds
- ✓ Keep your costs low
- ✓ Fully-managed autoscaling
- ✓ Enterprise security

Starting at
\$0.06 /hour


PRO Pro Account
Show your support for the best ML community

- ✓ Get a PRO badge on your profile
- ✓ Early access to new features
- ✓ Higher tier for the Free Inference API
- ✓ Higher tier for AutoTrain

Subscribe for
\$9 /month


AutoTrain
Create powerful AI models without code

- ✓ Automatic models search and training
- ✓ Easy drag and drop interface
- ✓ 9 tasks available (for Vision, NLP and more)
- ✓ Models instantly available on the Hub

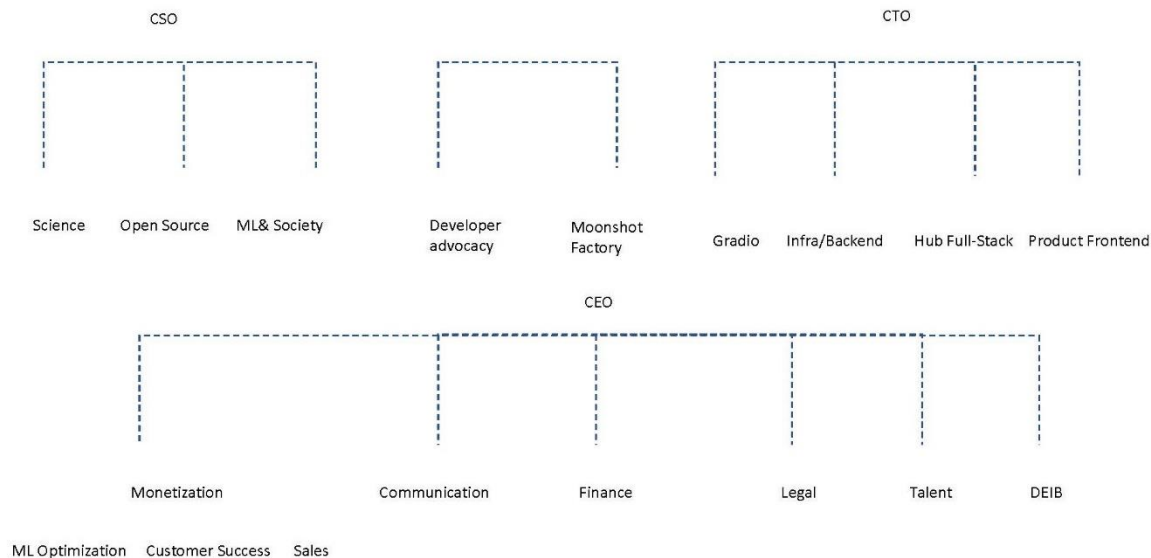
Starting at
\$0 /month


Enterprise
ML Solutions from research to production

- ✓ Accelerate your ML roadmap
- ✓ Direct support from our experts
- ✓ Private Hub deployment
- ✓ Hardware Partner Program

Contact us
Get Quote

Source: Company documents.

Exhibit 10a Simplified Hugging Face Organizational Structure, Fall 2022

A simplified view of the organizational structure of Hugging Face. The company was divided across priorities around Platform Development (CTO), Science (CSO), and Open-Source (also CSO), with the CEO overseeing operational and monetization functions that synthesize across these activities.

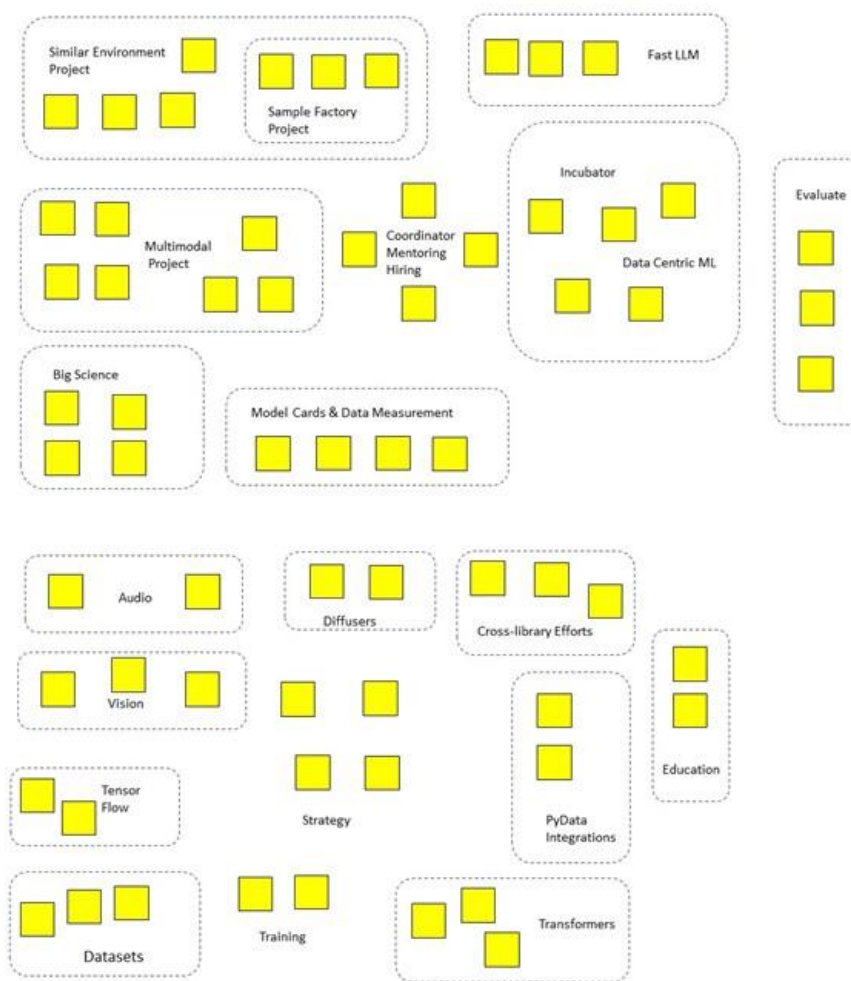
Under Julien Chaumond (CTO), the team pursued projects like supporting gradio (the technical library supporting ML demos, like Spaces), the product infrastructure (such as managing the technical interface with cloud compute resources), and, the Hugging Face website (including the website front-end and full-stack integration of the website with infrastructure).

Under Thomas Wolf (CSO), teams were organized into supporting novel scientific projects (such as developing large language models like BLOOM or simulation environments for training models), supporting or pursuing open-source development, and managing public policy and ethical considerations.

Finally, under Clément Delangue (CEO), teams were organized to support revenue-generating activities, such as sales and customer support. Delangue also oversaw a variety of administrative functions like communications, legal, and finance, which comprised one-person teams focused on helping others accomplish their respective functions.

Not every team fits neatly into this structure. For example, the developer advocacy team and moonshot factory were co-located under the CSO and CTO, and focused on interacting with the open-source developer community and trying out support for very specific exploratory verticals such as art or mobile support (respectively). Because teams were fluid and regularly reshuffled, the organizational chart was intended to add clarity but not restrict collaboration within the organization.

Source: Casewriters, based on company documents.

Exhibit 10b Organizational Chart for the Science and Open Source Teams, Fall 2022

Source: Company documents.

Exhibit 11 Funding Rounds (in US\$ millions)

Date	Round	Amount Raised	Lead Investor
March 5, 2017	Angel Round	1.2	Betaworks
May 23, 2018	Seed Round	4	Ronny Conway
December 17, 2019	Series A	15	Lux Capital
March 11, 2021	Series B	40	Addition
May 9, 2022	Series C	100	Lux Capital

Source: Crunchbase, "Hugging Face," undated, https://www.crunchbase.com/organization/hugging-face/company_financials, accessed October, 2022.

Exhibit 12 Hugging Face Organizational Priorities in 2022

Priority	Platform	Open Source	Scientific Research
Exemplar Projects	Add Enterprise MLOps Features Create open-source support for recent research around tabular data	Add support for non-NLP ML models like audio, vision, and, reinforcement learning	BigScience Collaboration to create BLOOM
Project Description	Develop deep learning approaches to leveraging data stored in rows and columns, like spreadsheets	Add library support for features that would make it easier to store, train, and deploy ML models focused on alternative data types	A new AI model that could translate many languages, as well as create its own text based on language prediction. BLOOM differed in that it was not generated by a single organization, but had many contributors, with Hugging Face working as an organizer of the many contributions. BLOOM was also supported by governments and other nonprofits, with the French government paying for the compute power necessary for training
Value to Community	Tabular data is common in organizations, and effective models here could have many applications in improving customer company operations	Quality support for alternative data types would enable larger parts of the AI community to benefit from Hugging Face's infrastructure	BLOOM is one of the first widely available Large Language model. Unlike other LLMs like GPT-3, its code was publicly available, enabling others to research the model and develop applications
Value to Hugging Face	Targeted feature development that would appeal to Hugging Face's paying customers	Broader support would attract more of the AI community onto the platform, expanding the technology offerings and increasing the amount of business use-cases supported by the platform	Successful research brings visibility to Hugging Face, increasing brand value and strengthening relationships with other researchers on the technical frontier

Source: Company documents.

Endnotes

- ¹ Mike Horia Teodorescu, “Machine Learning Methods for Strategy Research,” Harvard Business School working paper, October 14, 2017, <https://bit.ly/3DYXO68>, accessed October 2022.
- ² Teodorescu, “Machine Learning Methods for Strategy Research.”
- ³ Larry Hardesty, “Explained: Neural Networks,” MIT News website, April 14, 2017, <https://bit.ly/3zCNHkI>, accessed October 2022.
- ⁴ James Somers, “Is AI Riding a One-Trick Pony,” *Technology Review*, September 29, 2017, <https://bit.ly/3h2Kc0G>, accessed October 2022.
- ⁵ Dave Gershorn, “The data that transformed AI research – and possibly the world,” *Quartz*, July 26, 2017, <https://bit.ly/3TZngy7>, accessed October 2022.
- ⁶ Dave Gershorn, “The data that transformed AI research.”
- ⁷ Gershorn, “The data that transformed AI research.” For ImageNet, see <https://bit.ly/3T1Uo70>, accessed September 2018.
- ⁸ Wayne Duggan, “Artificial Intelligence Stocks: The 10 Best AI Companies,” U.S. News & World Report, September 9, 2022, <https://bit.ly/2u9N3w7>, accessed September 2022.
- ⁹ Lionel Sujay Vailshery, “Vendor market share in cloud infrastructure services market worldwide 2017-2022,” Statista, August 25, 2022, via Statista, accessed September 2022.
- ¹⁰ Russel Kay, “How to LAMP,” *Computerworld*, September 25, 2006, <https://bit.ly/3h0MYn3>, accessed October 2022.
- ¹¹ Frederic Lardinois and Ingrid Lunden, “Microsoft has acquired GitHub for \$7.5B in stock,” Tech Crunch, June 4, 2018, <https://techcrunch.com/2018/06/04/microsoft-has-acquired-github-for-7-5b-in-microsoft-stock/>, accessed September 2022.
- ¹² Ari Levy, “GitLab grew up in GitHub’s shadow – now it’s worth twice what Microsoft paid for its chief rival,” CNBC, October 17, 2021, <https://cnb.cx/3WpVBHY>, accessed August 2022.
- ¹³ Lardinois and Lunden, “Microsoft has acquired GitHub for \$7.5B in stock.”
- ¹⁴ Shane Greenstein and Frank Nagle, “Digital Dark Matter and the Economic Contribution of Apache,” *Research Policy*, 2014. V. 4, pp. 623-631, <https://doi.org/10.1016/j.respol.2014.01.003>, accessed November 2022.
- ¹⁵ Raviv Murciano-Goroff, Ran Zhuo, and Shane Greenstein, “Unsung Software and Veiled Value Creation: Illustrations from Server Software,” *Research Policy*, 2021, v. 50, pp. 1-31, <https://doi.org/10.1016/j.respol.2021.104333>, accessed November 2022.
- ¹⁶ Lionel Sujay Vailshery, “Open source services – worldwide revenue 2017-2022,” Statista, February 21, 2022, <https://bit.ly/3zzTAPR>, accessed September 2022.
- ¹⁷ Kenrick Cai, “The \$2 Billion Emoji: Hugging Face Wants To Be Launchpad For A Machine Learning Revolution,” *Forbes*, May 9, 2022, <https://bit.ly/3T4r0ge>, accessed August 2022.
- ¹⁸ Romain Dillet, “Hugging Face reaches \$2 Billion valuation to build the GitHub of machine learning,” TechCrunch, May 9, 2022, <https://tcrn.ch/3FOjSBA>, accessed October 2022.
- ¹⁹ Dillet, “Hugging Face raises \$4 million for its artificial BFF.”
- ²⁰ Crunchbase, “Hugging Face,” <https://bit.ly/3U39ASH>, accessed August 2022.
- ²¹ “Transformers Huge Community feedback: 40k,” Hugging Face, April 2021, <https://bit.ly/3FKVAJ3>, accessed May 2022.
- ²² “Julien Chaumond,” Dagshub Podcast, August 24, 2021, <https://bit.ly/3h6r1mK>, accessed October 2022.
- ²³ Antonis Papsavva, Savvas Zannettou, Emiliano De Cristofaro, Gianluca Stringhini, Jeremy Blackburn, “Dataset: Raiders of the Lost Kek: 3.5 Years of Augmented 4chan Posts from the Politically Incorrect Board,” Zenodo, January 13, 2020, <https://bit.ly/3WkZFJL>, accessed October 2022.
- ²⁴ Hugging Face, ykicher/gpt-4chan, no date, <https://huggingface.co/ykicher/gpt-4chan>, accessed October 2022.